# GENERALIZED NARAYANA ARRAYS, RESTRICTED DYCK PATHS, AND RELATED BIJECTIONS

JEAN-LUC BARIL, RIGOBERTO FLÓREZ, AND JOSÉ L. RAMÍREZ

Abstract. We study two types of generalizations of the Narayana array. Type 1 counts the number of Dyck paths of semilength $n$ with $k$ peaks (or valleys) whose heights belong to a given set $S$, whereas Type 2 counts Dyck paths of semilength $n$ with exactly $k$ peaks (or valleys), all of which have heights in $S$. We focus on the following sets: $S=\{m\}$, $S=\{m,m+1\}$, the non-negative even heights, the positive odd heights, and more generally heights congruent to a fixed residue modulo $a$. For each case, we provide a bivariate generating function $f(x,y)$ that characterizes the corresponding Narayana array, establish combinatorial connections through bijections with other combinatorial structures, and obtain asymptotic approximations for the expected number of peaks or valleys of height in $S$ in Dyck paths of a given semilength.

## 1. Introduction

A *Dyck path* is a lattice path in the first quadrant of the $xy$-plane that starts at the origin, ends on the $x$-axis, and consists of an equal number of up-steps $U=(1,1)$ and down-steps $D=(1,-1)$. The *length* of a Dyck path is defined as the total number of steps. For a Dyck path $P$, we write $|P|$ to denote its semilength, that is, half of its length. The empty Dyck path (a Dyck path of semilength 0) is denoted by $\epsilon$. We denote by $\mathcal{D}$ the set of all Dyck paths and by $\mathcal{D}_n$ the set of all Dyck paths of semilength $n$, for $n\geq 0$. It is well-known that the number of Dyck paths of semilength $n$ is given by the $n$-th Catalan number, $c_n=\frac{1}{n+1}\binom{2n}{n}$ (cf. [19]).

A *peak* in a Dyck path is a subpath of the form $UD$, while a *valley* is a subpath of the form $DU$. The *height* of a peak (resp. valley) is the $y$-coordinate of its highest point (resp. lowest point). Given $i\geq 0$ and a Dyck path $P$, we denote by $\texttt{pea}_i(P)$ (resp. $\texttt{val}_i(P)$) the number of peaks (resp. valleys) in $P$ at height $i$.

In this paper, we study two generalizations of the Narayana array to count Dyck paths of semilength $n$ whose peaks (or valleys) have heights in a fixed set $S$. We also consider the related problem of counting Dyck paths with exactly $k$ peaks (or valleys), all of which have heights in the set $S$. In particular, we examine special cases where $S$ is the set of non-negative even numbers, the set of positive odd numbers, $S=\{m\}$, $S=\{m,m+1\}$, and $S=\{m\in\mathbb{N} : m\equiv b \pmod a\}$. For each of these counting problems, we use a bivariate generating function $f(x,y)$ that characterizes the Narayana array. These

---

*Date*: April 29, 2025.
2010 *Mathematics Subject Classification.* 05A15, 05A19.
*Key words and phrases.* Dyck path, Narayana array, Catalan number, generating function.

generating functions give rise to two types of arrays, which we refer to as generalized Narayana arrays. Additionally, we present several bijective correspondences between these classes of Dyck paths and other families of paths, permutations, compositions or trees.

We derive the asymptotic behavior of the expected number of peaks (or valleys) at heights in $S$ among all Dyck paths of semilength $n$ by computing $\partial_y(f(x,y))|_{y=1}$ and applying classical methods (cf. [12]). We also extend the previous definitions of $\mathtt{pea}_i(P)$ and $\mathtt{val}_i(P)$ to define $\mathtt{pea}_S(P)$ (resp. $\mathtt{val}_S(P)$) when $i$ belongs to a given set $S$.

Finally, we include an appendix that presents a computational approach to derive these generating functions, based on recent work by Zeilberger [8, 10] and Bu [7] on restricted Dyck and Motzkin paths.

## 2. Some basic definitions, notation, and background

In this section, we introduce some basic definitions, notation, and background necessary for understanding the generalizations of the Narayana array and the related combinatorial problems studied in this paper.

Throughout the paper, we will often consider the *first return decomposition* of a nonempty Dyck path $R$, that is, the unique decomposition $R = UPDQ$, where $P$ and $Q$ are (possibly empty) Dyck paths. From this decomposition, we deduce the functional equation $C(x) = 1 + xC(x)^2$ for the generating function $C(x) = \sum_{n\geq 0} c_n x^n$, which implies that

$$C(x) = \frac{1-\sqrt{1-4x}}{2x}.$$

For further information on Catalan numbers, see A000108 in the On-line Encyclopedia of Integer Sequences [18]. Various works on the enumeration and generation of Dyck paths with respect to their semilength and other combinatorial statistics can be found in [1, 3, 9, 13, 14, 15, 17].

Let $N(n,k)$ denote the number of Dyck paths of semilength $n$ with exactly $k$ peaks. The array $(N(n,k))_{n,k\geq 0}$, known as the *Narayana triangle* (see [16] and A001263 in [18]), is given by the formula

$$N(n,k) = \frac{1}{k}\binom{n-1}{k-1}\binom{n}{k-1}, \quad n,k \geq 1.$$

Its generating function is

$$\sum_{n,k\geq 1} N(n,k)x^n y^k = \frac{1-x(y+1)-\sqrt{1-2x(y+1)+x^2(y-1)^2}}{2xy}.$$

Let $S$ be a subset of non-negative integers. Define $P_S(n,k)$ as the number of Dyck paths of semilength $n$ with exactly $k$ peaks whose heights belong to $S$. Note that the paths may also contain additional peaks whose heights are not in $S$. The array $\mathcal{P}_S^{(1)} := (P_S(n,k))_{n,k\geq 0}$ is called *Narayana array of type 1*. Clearly, whenever $S = \mathbb{N}$, the sequence $P_{\mathbb{N}}(n,k)$ coincides with the Narayana number $N(n,k)$.

Dually, let $V_S(n,k)$ denote the number of Dyck paths of semilength $n$ with exactly $k$ valleys whose heights are contained in $S$. The corresponding array, $\mathcal{V}_S^{(1)} := (V_S(n,k))_{n,k\geq 0}$, is called the *dual Narayana array of type 1.*

Eu, Liu, and Yeh [11] studied the sequence $(P_S(n,0))_n$, which counts the number of Dyck paths whose peak heights avoid $S$.

Another variation of the Narayana array imposes the restriction that the heights of all peaks or valleys must be contained in a given set. Specifically, let $\overline{P}_S(n,k)$ denote the number of Dyck paths of semilength $n$ with exactly $k$ peaks, where every peak has height in $S$. The array $\mathcal{P}_S^{(2)} := (\overline{P}_S(n,k))_{n,k\geq 0}$ is called the *Narayana array of type 2.*

Similarly, let $\overline{V}_S(n,k)$ denote the number of Dyck paths of semilength $n$ with exactly $k$ valleys, where every valley has height in $S$. The array $\mathcal{V}_S^{(2)} := (\overline{V}_S(n,k))_{n,k\geq 0}$ is called the *dual Narayana array of type 2.*

The notation is summarized in Table 1.

| | | | |
|---|---|---|---|
| $P_S(n,k)$ | number of Dyck paths of semilength $n$ with exactly $k$ peaks of heights in $S$ | $\mathcal{P}_S^{(1)}$ | Narayana array of type 1 |
| $V_S(n,k)$ | number of Dyck paths of semilength $n$ with exactly $k$ valleys of heights in $S$ | $\mathcal{V}_S^{(1)}$ | Dual Narayana array of type 1 |
| $\overline{P}_S(n,k)$ | number of Dyck paths of semilength $n$ with exactly $k$ peaks, each having a height in $S$ | $\mathcal{P}_S^{(2)}$ | Narayana array of type 2 |
| $\overline{V}_S(n,k)$ | number of Dyck paths of semilength $n$ with exactly $k$ valleys, each having a height in $S$ | $\mathcal{V}_S^{(2)}$ | Dual Narayana array of type 2 |

TABLE 1. Narayana sequences.

## 3. Narayana Arrays of Type 1. The Peak Statistic

In this section, we use generating functions to count Dyck paths with peaks at heights belonging to a given set $S$. We begin by analyzing the special case in which $S = \{m, m+1\}$, where the corresponding generating function can be expressed as a finite continued fraction. In particular, we describe a bijection that maps Dyck paths to Dyck paths with no peaks at heights 2 and 3. In the second part of the section, we examine the cases where $S$ consists of the positive even heights or the positive odd heights, showing that the resulting matrices are related to the number of permutations avoiding the pattern 132. Additionally, we consider sets $S$ defined by arithmetic progressions. Finally, we present corollaries describing the asymptotic behavior of Dyck paths based on the distribution of peak heights.

Let $\mathbb{N}$ and $\mathbb{Z}^+$ denote the set of non-negative integers and positive integers, respectively. Given $S \subset \mathbb{N}$, we define the sets $S_- := \{s-1 : s \in S\} \cap \mathbb{Z}^+$ and $S_-^* := \{s-1 : s \in S\} \cap \mathbb{N}$. Moreover, we define $[n] := \{1, 2, \ldots, n\}$. We introduce the following bivariate generating function:

$$\Delta_S(x,y) = \sum_{n,k\geq 0} P_S(n,k)x^n y^k,$$

where $P_S(n,k)$ is the number of all Dyck paths of semilength $n$ with $k$ peaks whose heights are contained in $S$.

3.1. **The statistics of peaks at heights in a given set.** In this section, we count the number of Dyck paths with peaks at heights in set $S$.

**Theorem 3.1.** *The number of Dyck paths with peaks at heights in $S \subseteq \mathbb{Z}^+$ is counted by the generating function $\Delta_S(x,y)$, which satisfies the following equation:*

$$\Delta_S(x,y) = \begin{cases} 1 + xy\Delta_S(x,y) + x(\Delta_{S_-}(x,y) - 1)\Delta_S(x,y), & \text{if } 1 \in S; \\ 1 + x\Delta_{S_-}(x,y)\Delta_S(x,y), & \text{otherwise.} \end{cases}$$

*Moreover, $\Delta_\emptyset(x,y) = C(x)$, where $C(x)$ is the generating function of the Catalan numbers.*

*Proof.* Let $R$ be a non-empty Dyck path, and let $R = UPDQ$ be its first return decomposition, where $P$ and $Q$ are possibly empty Dyck paths. If $1 \in S$, then we distinguish two separate cases: either $R = UDQ$ or $R = UPDQ$, where $P$ and $Q$ are Dyck paths, with $P$ being non-empty.

For the first case, $R$ has one more peak at height $1 \in S$ than $Q$, the contribution $xy\Delta_S(x,y)$. In the second case, $R$ has $a+b$ peaks at heights in $S$, where $a$ (resp. $b$) is the number of peaks in $P$ (resp. $Q$) at heights in $S_-$ (resp. $S$). Considering these two cases, we obtain the functional equation:

$$\Delta_S(x,y) = 1 + xy\Delta_S(x,y) + x(\Delta_{S_-}(x,y) - 1)\Delta_S(x,y).$$

If $1 \notin S$, then the decomposition $R = UPDQ$ where $P$ and $Q$ are Dyck paths, leads the functional equation

$$\Delta_S(x,y) = 1 + x\Delta_{S_-}(x,y)\Delta_S(x,y).$$

Notice that if $S = \emptyset$ in this last equation, we recover the functional equation for Dyck paths:

$$\Delta_S(x,y) = 1 + x\Delta_S(x,y)^2,$$

which implies that $\Delta_\emptyset(x,y) = C(x)$, where $C(x)$ is the generating function for Catalan numbers. □

In [9], Deutsch proved that the number of Dyck paths with no peaks at height 1 is counted by the Fine numbers A000957, while the number of Dyck paths with no peaks at height 2 are counted by the Catalan numbers. These two sequences are the first column of the two arrays $\mathcal{P}^{(1)}_{\{1\}}$ and $\mathcal{P}^{(1)}_{\{2\}}$.

3.1.1. *The statistics of peaks at heights in $S = \{m, m+1\}$.* In this section, we study the special case in which $S = \{m, m+1\}$. We also prove that Dyck paths with no peaks in $S = \{2, 3\}$ are counted by a shift of the Catalan sequence.

The proof of the following result is straightforward using Theorem 3.1.

**Corollary 3.2.** *The number of Dyck paths with peaks at heights in $\{m, m+1\}$ is counted by the generating function $\Delta_{\{m,m+1\}}(x, y)$, given by the continued fraction*

$$\Delta_{\{m,m+1\}}(x,y) = \left.\cfrac{1}{1-\cfrac{x}{1-\cfrac{x}{\cfrac{\vdots}{1-\cfrac{x}{1-xy+x-\cfrac{x}{1-xy-x(C(x)-1)}}}}}}\right\} m \text{ times},$$

*where $C(x)$ is the generating function for Catalan numbers.*

For instance, setting $m = 2$ in Corollary 3.2, so that $S = \{2, 3\}$, we obtain the generating function $\Delta_{\{2,3\}}(x, y)$:

$$\frac{2+3x-2(1+4x+3x^2+x^3)y+6x(1+x)^2y^2-6x^2(1+x)y^3+2x^3y^4-x\sqrt{1-4x}}{2(1-(1+3x)y+x(3+4x+x^2)y^2-x^2(3+2x)y^3+x^3y^4)}.$$

From this generating function, we can obtain the first few rows of the Narayana array of type 1 associated with the set $S = \{2, 3\}$.

$$\mathcal{P}^{(1)}_{\{2,3\}} = \begin{pmatrix} 1 & 0 & 0 & 0 & 0 & 0 & 0 \\ 1 & 0 & 0 & 0 & 0 & 0 & 0 \\ 1 & 1 & 0 & 0 & 0 & 0 & 0 \\ 1 & 3 & 1 & 0 & 0 & 0 & 0 \\ 2 & 5 & \boxed{\mathbf{6}} & 1 & 0 & 0 & 0 \\ 5 & 11 & 15 & 10 & 1 & 0 & 0 \\ 14 & 29 & 38 & 35 & 15 & 1 & 0 \\ 42 & 84 & 107 & 104 & 70 & 21 & 1 \end{pmatrix}.$$

For example, $P^{(1)}_{\{2,3\}}(4, 2) = 6$, which corresponds to the boxed entry in the array $\mathcal{P}^{(1)}_{\{2,3\}}$ above. The six Dyck paths corresponding to this value are illustrated in Figure 1. The red points denote the peaks of height 2 or 3.

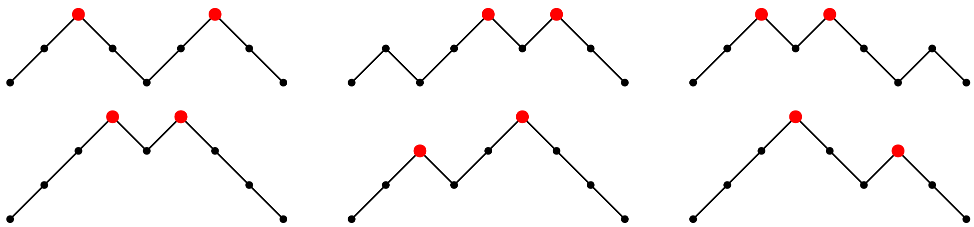


FIGURE 1. Dyck paths of semilength 4 with 2 peaks of heights 2 or 3.

The first column of $\mathcal{P}^{(1)}_{\{2,3\}}$, consisting of Dyck paths with no peaks in $S = \{2, 3\}$, is a shifted version of the Catalan sequence A000108. Let us define a recursively map $\chi$ from Dyck paths to Dyck paths with no peaks at heights 2 or 3:

$$\chi(\epsilon) = UDUD,$$

$$\chi(UDP) = UD\chi(P),$$

$$\chi(P_1(UD)^{a_1}P_2(UD)^{a_2}\cdots(UD)^{a_{k-1}}P_k(UD)^{a_k}) = UUP_1f(a_1)P_2f(a_2)\cdots f(a_{k-1})P_kDD(UD)^{a_k},$$

where $k \geq 1$, $P_i = UQ_iD$ with $Q_i \neq \epsilon$ for $1 \leq i \leq k$, and $a_i \geq 1$ for $1 \leq i \leq k-1$, $a_k \geq 0$, and $f(1) = DU$, $f(a) = DD(UD)^{a-2}UU$ for $a \geq 2$.

For example, if

$$R = UUDUDDUDUUDDUUDDUDUDUUDDUDUDUUDDUUDUDD,$$

then we have

$$\chi(R) = UUUUDUDDDUUUDDUUDDDDUUUUDDDDUUUUDDUUDUDDDD.$$

Figure 2, illustrates the Dyck paths corresponding to this example.

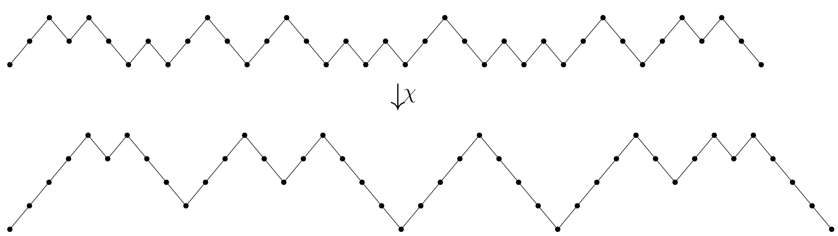



FIGURE 2. Example of the bijection $\chi$.

By induction on the semilength, the image by $\chi$ of a Dyck path of semilength $n$ is a Dyck path of semilength $n + 2$ with no peaks at height 2 or 3, and by simple observation $\chi$ is injective. Furthermore, the generating function corresponding to the first column of $\mathcal{P}^{(1)}_{\{2,3\}}$ satisfies $\Delta_{\{2,3\}}(x, 0) = 1 + x + x^2C(x)$, which implies the bijectivity of $\chi$ between Dyck paths of semilength $n$ and Dyck paths of semilength $n + 2$ with no peaks at height 2 or 3, thus offering an alternative combinatorial interpretation of the Catalan numbers.

By calculating $\partial_y(\Delta(x, y))|_{y=1}$, and applying classical methods (see [12]), we deduce the following result.

**Corollary 3.3.** *An asymptotic estimate for the expected number of peaks at heights in* $\{2, 3\}$ *among all Dyck paths of semilength* $n$ *is* 5.

3.1.2. *The statistics of peaks at either even or odd heights.* In this section, we present generating functions that count the number of Dyck paths with peaks at even heights and at odd heights, respectively. We use $\mathbb{E}$ to denote the set of all non-negative even integers and $\mathbb{O}$ to denote the set of all positive odd integers.

**Corollary 3.4.** *The number of Dyck paths with peaks at even heights is counted by the generating function* $\Delta_{\mathbb{E}}(x, y)$*, which satisfies the following equation:*

$$\Delta_{\mathbb{E}}(x, y) = \frac{1 + x - xy - \sqrt{1 - 2x(1 + y) + x^2(-3 + 2y + y^2)}}{2x}.$$

*Similarly, the number of Dyck paths with peaks at odd heights is counted by the generating function* $\Delta_{\mathbb{O}}(x, y)$*, which satisfies the following equation:*

$$\Delta_{\mathbb{O}}(x, y) = \frac{1 + x - xy - \sqrt{1 - 2x(1 + y) + x^2(-3 + 2y + y^2)}}{2x(1 + x(1 - y))}.$$

The proof of this corollary follows by solving the system of equations derived from Theorem 3.1.

$$\begin{cases} \Delta_{\mathbb{E}}(x, y) = & 1 + x\Delta_{\mathbb{O}}(x, y)\Delta_{\mathbb{E}}(x, y), \\ \Delta_{\mathbb{O}}(x, y) = & 1 + x(\Delta_{\mathbb{E}}(x, y) - 1)\Delta_{\mathbb{O}}(x, y) + xy\Delta_{\mathbb{O}}(x, y). \end{cases}$$

From Corollary 3.4, we obtain the first few rows of the Narayana arrays of type 1 associated with the two sets of even and odd integers.

$$\mathcal{P}_{\mathbb{E}}^{(1)} = \begin{pmatrix} 1 & 0 & 0 & 0 & 0 & 0 & 0 \\ 1 & 0 & 0 & 0 & 0 & 0 & 0 \\ 1 & 1 & 0 & 0 & 0 & 0 & 0 \\ 2 & 2 & 1 & 0 & 0 & 0 & 0 \\ 4 & 6 & 3 & 1 & 0 & 0 & 0 \\ 9 & 16 & 12 & 4 & 1 & 0 & 0 \\ 21 & 45 & 40 & 20 & 5 & 1 & 0 \\ 51 & 126 & 135 & 80 & 30 & 6 & 1 \end{pmatrix} \quad \text{and} \quad \mathcal{P}_{\mathbb{O}}^{(1)} = \begin{pmatrix} 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 \\ 1 & 0 & 1 & 0 & 0 & 0 & 0 & 0 \\ 1 & 3 & 0 & 1 & 0 & 0 & 0 & 0 \\ 3 & 4 & 6 & 0 & 1 & 0 & 0 & 0 \\ 6 & 15 & 10 & 10 & 0 & 1 & 0 & 0 \\ 15 & 36 & 45 & 20 & 15 & 0 & 1 & 0 \\ 36 & 105 & 126 & 105 & 35 & 21 & 0 & 1 \end{pmatrix}.$$

The array $\mathcal{P}_{\mathbb{E}}^{(1)}$ coincides with Table 2 given by Baril and Kirgizov in [2], where the entry of row $n$ and column $k$ counts the number of permutations of length $n$ avoiding the pattern 132 and having $k$ pure descents. A fundamental notion in the study of permutation patterns is that a permutation $\pi = \pi_1\pi_2 \cdots \pi_n$ *avoids* the pattern 132 if there does not exist $i_1 < i_2 < i_3$ such that $\pi_{i_1} < \pi_{i_3} < \pi_{i_2}$. A *pure descent* in a permutation $\pi$ is an occurrence $\pi_i > \pi_{i+1}$ where there is no $j < i$ such that $\pi_i < \pi_j < \pi_{i+1}$. From [2], we can directly deduce that Dyck paths of a given semilength with no peak at even heights are counted by Motzkin numbers (see A001006).

We now define a recursive map $\phi$ from the set of Dyck paths to the set of 132-avoiding permutations as follows: $\phi(\epsilon)=\epsilon$, $\phi(UDP)=\phi(P)\ n$ with $n=1+|P|$, and

$$\begin{aligned}\phi(UUQ_1D\cdots UQ_kDDP) = &\\ &(\phi(Q_1)+n_2+\cdots+n_k+m+k)(n_1+\cdots+n_k+m+k+1)\\ &(\phi(Q_2)+n_3+\cdots+n_k+m+k-1)(n_2+\cdots+n_k+m+k)\\ &\cdots\\ &(\phi(Q_{k-1})+n_k+m+2)(n_{k-1}+n_k+m+3)\\ &(\phi(Q_k)+m+1)(n_k+m+2)\\ &\phi(P)(m+1),\end{aligned}$$

where $Q_1, Q_2, \ldots, Q_k, P$ are some Dyck paths with $n_i=|Q_i|$ for $1\le i\le k$, and $m=|P|$.

By induction on the semilength, we establish that the image of a Dyck path under $\phi$ is a permutation avoiding the pattern 132, Furthermore, it follows (again by induction) that $\phi$ is injective. Since the number of permutations avoiding 132 of a given length are counted by the Catalan numbers, which also enumerate Dyck paths, we deduce that $\phi$ is a bijection. Moreover, an induction on the semilength shows that the number of peaks at even heights in a Dyck path $R$ equals the number of pure descents in the permutation $\phi(R)$. This claim is immediate for $R=UDP$. For $R=UUQ_1D\cdots UQ_kDDP$, $\texttt{pea}_{\mathbb{E}}(R)$ is the number of empty paths $Q_i$ plus

$$\texttt{pea}_{\mathbb{E}}(Q_1)+\texttt{pea}_{\mathbb{E}}(Q_2)+\cdots+\texttt{pea}_{\mathbb{E}}(Q_k)+\texttt{pea}_{\mathbb{E}}(P).$$

By the induction hypothesis, this sum also counts the number of empty paths $Q_i$ plus the number of pure descents in all $Q_i$ and $P$. Since an empty path $Q_i$ generates a pure descent of the form $(n_{i-1}+\cdots+n_k+m+k+1)a$, where $a$ is the first value of $(\phi(Q_i)+n_{i+1}+\cdots+n_k+m+k-1)$, it follows that $\texttt{pea}_{\mathbb{E}}(R)$ matches the number of pure descents in $\phi(R)$, completing the induction.

For instance, if $R = U\ UD\ UUDUDD\ UD\ D\ UUDD$, then we obtain $\phi(R) =$ 8 567 4 21 3, and $\texttt{pea}_{\mathbb{E}}(R)=3$, which is also the number of pure descents in $\phi(R)$ (namely, 85, 42, and 21), see Figure 3.

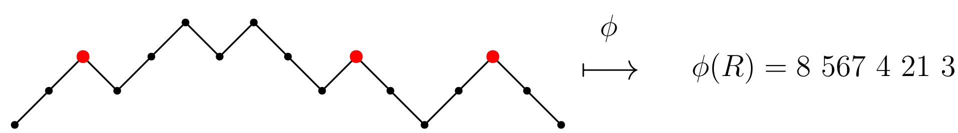



FIGURE 3. Example of the bijection $\chi$.

Moreover, the bijection $\phi$ maps the number $\texttt{pea}_{\mathbb{O}}(R)$ of peaks at odd heights in a Dyck path $R$ of semilength $n$, to the number of adjacencies in the extended permutation $0\ \phi(R)(n+1)$, where an adjacency is an occurrence of two consecutive entries in the form $a(a+1)$. In the above example, we have $\texttt{pea}_{\mathbb{O}}(R)=2$ and the permutation $0\ \phi(R)\ 9$ has two adjacencies 56 and 67.

**Corollary 3.5.** *An asymptotic estimate for the expected number of peaks at even heights (resp. odd heights) among all Dyck paths of semilength $n$ is $n/4$ (resp. $n/4$).*

3.1.3. *The statistics of peaks at heights in arithmetical progressions.* In this section, we generalize the above result to sets defined by an arithmetic progression. As before, we provide generating functions that enumerate Dyck paths whose peaks occur at heights within these sets.

Let $a$ and $b$ be integers, with $0 \leq b < a$. We define $A_{a,b} = \{m \in \mathbb{N} : m \equiv b \pmod{a}\}$ as the set of residue classes modulo $a$.

**Corollary 3.6.** *The number of Dyck paths with peaks at $A_{3,0}$, $A_{3,1}$, and $A_{3,2}$ is counted by generating functions satisfying the following equations:*

$$\Delta_{A_{3,0}}(x,y) = \frac{1 - x^2 - xy + x^2y - \sqrt{(1-x)((1-x)(1+x(1-y))^2 - 4x(1-xy))}}{2(1-x)x},$$

$$\Delta_{A_{3,1}}(x,y) = \frac{1-x}{1-xy}\Delta_{A_{3,0}}(x,y),$$

$$\Delta_{A_{3,2}}(x,y) = \frac{1}{1-xy}\left((1-x)\Delta_{A_{3,0}}(x,y) + x(1-y)\right).$$

The proof of this corollary follows by solving the system of equations derived from Theorem 3.1.

$$\begin{cases} \Delta_{A_{a,0}}(x,y) = 1 + x\Delta_{A_{a,a-1}}(x,y)\Delta_{A_{a,0}}(x,y), \\ \Delta_{A_{a,1}}(x,y) = 1 + x(\Delta_{A_{a,0}}(x,y) - 1)\Delta_{A_{a,1}}(x,y) + xy\Delta_{A_{a,1}}(x,y), \\ \vdots \\ \Delta_{A_{a,a-1}}(x,y) = 1 + x\Delta_{A_{a,a-2}}(x,y)\Delta_{A_{a,a-1}}(x,y). \end{cases}$$

From Corollary 3.6, we obtain the first few rows of the corresponding Narayana arrays.

$$\mathcal{P}^{(1)}_{A_{3,0}} = \begin{pmatrix} 1 & 0 & 0 & 0 & 0 & 0 \\ 1 & 0 & 0 & 0 & 0 & 0 \\ 2 & 0 & 0 & 0 & 0 & 0 \\ 4 & 1 & 0 & 0 & 0 & 0 \\ 9 & 4 & 1 & 0 & 0 & 0 \\ 22 & 14 & 5 & 1 & 0 & 0 \\ 57 & 47 & 21 & 6 & 1 & 0 \\ 154 & 155 & 83 & 29 & 7 & 1 \end{pmatrix}, \quad \mathcal{P}^{(1)}_{A_{3,1}} = \begin{pmatrix} 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 \\ 1 & 0 & 1 & 0 & 0 & 0 & 0 & 0 \\ 2 & 2 & 0 & 1 & 0 & 0 & 0 & 0 \\ 5 & 5 & 3 & 0 & 1 & 0 & 0 & 0 \\ 13 & 15 & 9 & 4 & 0 & 1 & 0 & 0 \\ 35 & 46 & 31 & 14 & 5 & 0 & 1 & 0 \\ 97 & 143 & 108 & 54 & 20 & 6 & 0 & 1 \end{pmatrix},$$

$$\mathcal{P}^{(1)}_{A_{3,2}} = \begin{pmatrix} 1 & 0 & 0 & 0 & 0 & 0 & 0 \\ 1 & 0 & 0 & 0 & 0 & 0 & 0 \\ 1 & 1 & 0 & 0 & 0 & 0 & 0 \\ 2 & 2 & 1 & 0 & 0 & 0 & 0 \\ 5 & 5 & 3 & 1 & 0 & 0 & 0 \\ 13 & 15 & 9 & 4 & 1 & 0 & 0 \\ 35 & 46 & 31 & 14 & 5 & 1 & 0 \\ 97 & 143 & 108 & 54 & 20 & 6 & 1 \end{pmatrix}.$$

The first column of $\mathcal{P}^{(1)}_{A_{3,0}}$ corresponds to the sequence A105633, where the $n$-th term is the number of Dyck paths of semilength $n$ that avoid the pattern $UUDU$.

We now define a recursive map $\mu$ from the set of Dyck paths avoiding $UUDU$ to the set of Dyck paths with no peaks at heights congruent to 0 modulo 3, as follows:

$$\begin{aligned}\mu(\epsilon) &= \epsilon,\\ \mu(UDP) &= UD\mu(P),\\ \mu(UUDDP) &= UUDD\mu(P),\\ \mu(U^{k+2}DDQ_1DQ_2D\cdots Q_kDP) &= UUf(Q_1)f(Q_2)\cdots f(Q_k)DD\mu(P),\end{aligned}$$

where $k \geq 1$, $Q_1, \dots, Q_k, P$ are Dyck paths, $f(\epsilon) = DU$, and $f(P) = U\mu(P)D$ when $P \neq \epsilon$.

By a simple induction on the semilength, we can show that $\mu(R)$ is a Dyck path with no peaks at heights congruent to 0 modulo 3 whenever $R$ avoids $UUDU$ and that $\mu$ is injective. Since both classes of Dyck paths have the same cardinality, it follows that $\mu$ is a bijection.

For instance, if $R = UDUUUDDUDUUUDDUDDDUUDD$, then we obtain

$$\mu(R) = UDUUUUDUUUUDDDDDDDDUUDD,$$

and $\mu(R)$ has no peaks at heights congruent to 0 modulo 3. See, for example, Figure 4.

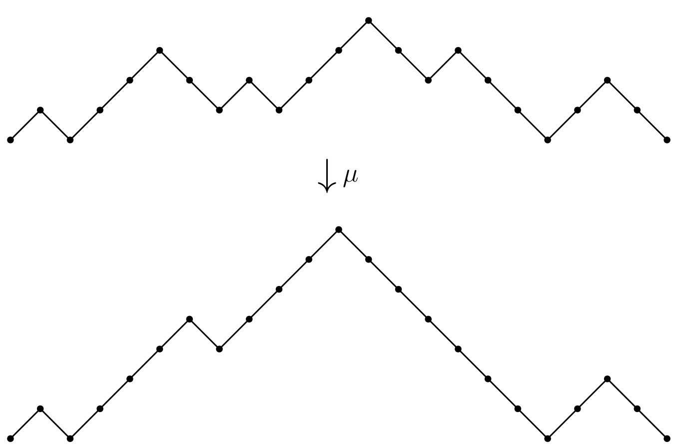


FIGURE 4. Example of the bijection $\mu$.

By observing the arrays $\mathcal{P}^{(1)}_{A_{3,2}}$ and $\mathcal{P}^{(1)}_{A_{3,1}}$, we believe that there is a bijection $\mu'$ on Dyck paths that maps the number of peaks at heights congruent to 2 modulo 3 to the number of peaks at heights congruent to 1 modulo 3 except for the path $U(UD)^{n-1}D$, whose image will be $(UD)^n$. Below, we describe such a bijection $\mu'$:

$$\begin{aligned}
\mu'(U(UD)^{n-1}D) &= (UD)^n \\
\mu'(UUPDQDR) &= UURDPDQ \text{ if } P \neq \epsilon \\
\mu'(U(UD)^k DUQDR) &= (UD)^k UURDDQ \text{ if } k \geq 1 \\
\mu'(U(UD)^k UPDQDR) &= (UD)^k UURDPDQ \text{ if } k \geq 1 \text{ and } P \neq \epsilon \\
\mu'(UDUQDR) &= UURDDQ,
\end{aligned}$$

where $P, Q, R$ are some Dyck paths, the five cases in the definition are distinct and form a partition of the set of Dyck paths. Consequently, the map $\mu'$ is injective, and by a cardinality argument, it is also bijective. Moreover, $\mu'$ preserves the relevant statistics as expected, except for the special Dyck path $U(UD)^{n-1}D)$.

For instance, there are five Dyck paths of semilength 4 with exactly one peak at height congruent to 2 modulo 3, namely
$UDUDUUDD$, $UDUUDDUD$, $UUDDUDUD$, $UUDUUDDD$, $UUUDDUDD$, and their images under $\mu'$ are $UUUUDDDD$, $UUUDDDUD$, $UDUUUDDD$, $UDUUDUDD$ and $UUDUDDUD$, respectively.

Moreover, the first column of $\mathcal{P}^{(2)}_{A_{3,2}}$ corresponds to the sequence A082582, where the $n$-th term represents the number of Dyck paths of semilength $n$ that avoid the pattern $UUDD$. We define the map $\mu''$ recursively from Dyck paths avoiding $UUDD$ to Dyck paths with no peak at height congruent to 2 modulo 3 as follows:

$$\begin{aligned}
\mu''(\epsilon) &= \epsilon \\
\mu''(UPDQ) &= U\mu'(\mu''(P))D\mu''(Q), \text{ if } P \neq UD.
\end{aligned}$$

Due to this recursive definition and the fact that $\mu'$ is a bijection between the set of Dyck paths avoiding peaks at height congruent to 2 modulo 3 and the set of Dyck paths avoiding peaks at height congruent to 1 modulo 3 (for Dyck paths of semilength at most 2), we can easily verify that $\mu''$ is a bijection mapping a Dyck path avoiding $UUDD$ to the set of Dyck paths avoiding peaks at height congruent to 2 modulo 3. For instance, we have

$$\begin{aligned}
\mu''(UDUUDUDD) &= UD\mu''(UUDUDD) = UDU\mu'(\mu''(UDUD))D \\
&= UDU\mu'(UDUD)D = UDUUUDDD.
\end{aligned}$$

**Corollary 3.7.** *An asymptotic estimate for the expected number of peaks of heights congruent to $i$ modulo 3 ($0 \leq i \leq 2$) among all Dyck paths of semilength $n$ is given by $n/6$ (for each $i = 0, 1, 2$).*

## 4. Narayana Arrays of Type 1. The Valley Statistic

In this section, we use generating functions that count Dyck paths with $k$ valleys belonging to a given set $S$. We first focus on the case $S = \{m, m+1\}$, and exhibit a simple bijection that relates valleys at heights $m$ and $m+1$ to peaks at heights $m+2$ and $m+3$. For other choices of $S$, such as the set of positive even integers or the set of positive odd integers, we construct several bijections. For example, we provide a bijection on Dyck paths that transforms the number of peaks of even height into the number of valleys of even height, and another bijection that relates the number of valleys at odd heights to the number of occurrences of the pattern $UUU$. We also obtain some results when $S$ is an arithmetical progression. Finally, we describe the asymptotic behavior of the corresponding counting sequences.

We introduce the bivariate generating function

$$\nabla_S(x,y) = \sum_{n,k\geq 0} V_S(n,k)x^n y^k,$$

where $V_S(n,k)$ is the number of all Dyck paths of semilength $n$ with $k$ valleys whose heights are contained in $S$.

Using arguments similar to those in Theorem 3.1, we derive the generating function with respect to the semilength and the number of valleys in $S$; therefore, we omit the proof.

**Theorem 4.1.** *The number of Dyck paths with valleys at heights in $S \subseteq \mathbb{N}$ is counted by the generating function $\nabla_S(x,y)$, which satisfies the following equation:*

$$\nabla_S(x,y) = \begin{cases} 1 + x\nabla_{S^*_-}(x,y)\nabla_S(x,y), & \text{if } 0 \notin S; \\ 1 + x\nabla_{S^*_-}(x,y) + xy\nabla_{S^*_-}(x,y)(\nabla_S(x,y) - 1), & \text{otherwise.} \end{cases}$$

*Moreover, $\nabla_\emptyset(x,y) = C(x)$, where $C(x)$ is the generating function of the Catalan numbers.*

4.0.1. *The statistics of valleys at heights in $S = \{m, m+1\}$.* In this section, we study the special case in which $S = \{m, m+1\}$ and $m \geq 0$.

The proof of the following corollary is a straightforward application of Theorem 4.1 and Theorem 3.1; therefore, we omit the proof.

**Corollary 4.2.** *The number of Dyck paths with valleyes at heights in $\{m, m+1\}$ is counted by the generating function $\nabla_{\{m,m+1\}}(x,y)$, which satisfies the following equation:*

$$\nabla_{\{m,m+1\}}(x,y) = \Delta_{\{m+2,m+3\}}(x,y),$$

*where $m \geq 0$.*

There exists a simple bijection on Dyck paths that maps the number of valleys at heights in $\{m, m+1\}$, $m \geq 0$, to the number of peaks at heights in $\{m+2, m+3\}$.

The transformation is as follows: reading a Dyck path $R$ from left to right, if we find a peak $UD$ at height in $\{m+2, m+3\}$, we replace it with $DU$, creating a valley at height in $\{m, m+1\}$. Conversely, if we find a valley at height in $\{m, m+1\}$, we replace it with $UD$, creating a peak at height in $\{m+2, m+3\}$.

For example, setting $m = 1$ and taking $R = UUDUUDDDUDUUDUUDDUDD$, we obtain the transformed path $UUUDDUDDUDUUUDDUUDDD$.

Moreover, in the case $m = 1$, that is $S = \{1, 2\}$, we obtain the generating function

$$\nabla_{\{1,2\}}(x,y) = \Delta_{\{3,4\}}(x,y).$$

The first few rows of the associated Narayana array of type 1 are

$$\mathcal{V}^{(1)}_{\{1,2\}} = \begin{pmatrix} 1 & 0 & 0 & 0 & 0 & 0 & 0 \\ 1 & 0 & 0 & 0 & 0 & 0 & 0 \\ 2 & 0 & 0 & 0 & 0 & 0 & 0 \\ 4 & 1 & 0 & 0 & 0 & 0 & 0 \\ 8 & 5 & \boxed{\mathbf{1}} & 0 & 0 & 0 & 0 \\ 17 & 16 & 8 & 1 & 0 & 0 & 0 \\ 38 & 48 & 33 & 12 & 1 & 0 & 0 \\ 90 & 140 & 119 & 62 & 17 & 1 & 0 \\ 227 & 407 & 404 & 260 & 108 & 23 & 1 \end{pmatrix}.$$

For example, $\mathcal{V}^{(1)}_{\{1,2\}}(4,2) = 1$, which corresponds to the boxed entry in the array $\mathcal{V}^{(1)}_{\{1,2\}}$ above. The Dyck path corresponding to this value is $UUDUDUDD$. This array is not listed in the OEIS.

**Corollary 4.3.** *An asymptotic estimate for the expected number of valleys at heights in* $\{0,1\}$ *among all Dyck paths of semilength* $n$ *is* 5.

4.0.2. *The statistics of valleys at either even or odd heights.* In this section, we provide generating functions that count Dyck paths according to the number of valleys at even heights and the number of valleys at odd heights, respectively.

**Corollary 4.4.** *We have*

$$\begin{aligned} \nabla_{\mathbb{E}}(x,y) =& \frac{1+x-xy-\sqrt{x^2y^2+2\,x^2y-3\,x^2-2\,xy-2\,x+1}}{2x} \quad \textit{and} \\ \nabla_{\mathbb{O}}(x,y) =& \frac{-xy+x-1+\sqrt{x^2y^2+2\,x^2y-3\,x^2-2\,xy-2\,x+1}}{2x\,(xy-x-y)}. \end{aligned}$$

The proof of this corollary follows by solving the system of equations derived from Theorem 4.1.

$$\begin{cases} \nabla_{\mathbb{E}}(x,y) = & 1 + x\nabla_{\mathbb{O}}(x,y)\nabla_{\mathbb{E}}(x,y), \\ \nabla_{\mathbb{O}}(x,y) = & 1 + x(\nabla_{\mathbb{E}}(x,y)-1)\nabla_{\mathbb{O}}(x,y) + xy\nabla_{\mathbb{O}}(x,y). \end{cases}$$

From Corollary 3.4, we obtain the first few rows of the Narayana arrays of type 1, $\mathcal{V}^{(1)}_{\mathbb{E}}$ and $\mathcal{V}^{(1)}_{\mathbb{O}}$, associated with the two sets of even and odd integers.

$$\mathcal{V}^{(1)}_{\mathbb{E}} = \begin{pmatrix} 1 & 0 & 0 & 0 & 0 & 0 & 0 \\ 1 & 0 & 0 & 0 & 0 & 0 & 0 \\ 1 & 1 & 0 & 0 & 0 & 0 & 0 \\ 2 & 2 & 1 & 0 & 0 & 0 & 0 \\ 4 & 6 & 3 & 1 & 0 & 0 & 0 \\ 9 & 16 & 12 & 4 & 1 & 0 & 0 \\ 21 & 45 & 40 & 20 & 5 & 1 & 0 \\ 51 & 126 & 135 & 80 & 30 & 6 & 1 \end{pmatrix} \quad \text{and} \quad \mathcal{V}^{(1)}_{\mathbb{O}} = \begin{pmatrix} 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 2 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 4 & 1 & 0 & 0 & 0 & 0 & 0 & 0 \\ 9 & 4 & 1 & 0 & 0 & 0 & 0 & 0 \\ 21 & 15 & 5 & 1 & 0 & 0 & 0 & 0 \\ 51 & 50 & 24 & 6 & 1 & 0 & 0 & 0 \\ 127 & 161 & 98 & 35 & 7 & 1 & 0 & 0 \end{pmatrix}.$$

Notice that $\mathcal{V}_{\mathbb{E}}^{(1)} = \mathcal{P}_{\mathbb{E}}^{(1)}$, and this equality can be verified by constructing the following bijection on Dyck paths. Define the map $\chi$ recursively to transform Dyck paths as follows:

$$\begin{aligned}
\chi(\epsilon) &= \epsilon,\\
\chi(UD) &= UD,\\
\chi(UDP) &= UU\underline{\chi(P)}, \text{ if } P \neq \epsilon\\
\chi(UUQ_1D\cdots UQ_kDD) &= Uf(UQ_1D)f(UQ_2D)\cdots f(UQ_kD)D\\
\chi(UUQ_1D\cdots UQ_kDDP) &= Uf(UQ_1D)f(UQ_2D)\cdots f(UQ_kD)U\underline{\chi(P)}, \text{ if } P \neq \epsilon.
\end{aligned}$$

Here, $\underline{P}$ (for $P \neq \epsilon$), is obtained from $P$ by replacing the first step $U$ with $D$. The function $f$ is defined by $f(UD) = DU$ and $f(UPD) = U\chi(P)D$ when $P \neq \epsilon$.

Using induction on the semilength, we can easily prove that $\chi$ is a bijection on Dyck paths that transforms the number of peaks of even height into the number of valleys of even height. More explicitly, $\chi(R)$ is obtained from $R$ by the following process: reading $R$ from left to right, if we find a peak $UD$ of even height, we replace it with $DU$ (creating a valley of even height); similarly, if we find a valley $DU$ of even height, we replace it with $UD$ (creating a peak of even height).

For instance, if $R = UUD\ UUDUDD\ UD\ D\ UUDD$, then we have

$$\chi(R) = UDUUUUDDDDUUDDUD.$$

The number of peaks of even height in $R$ is three, which is also the number of valleys of even height in $\chi(R)$; see Figure 5.

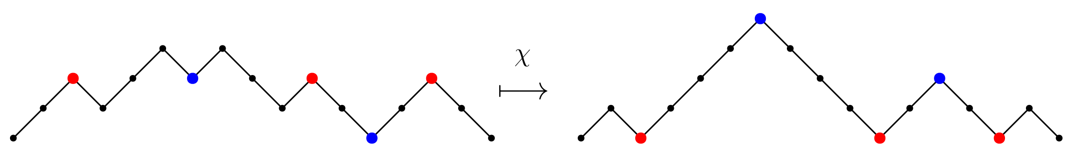



FIGURE 5. Example of the bijection $\chi$.

In addition, the array $\mathcal{V}_{\mathbb{O}}^{(1)}$ coincides with the array A092107, where the entry $V(n,k)$ represents the number of Dyck paths of semilength $n$ containing exactly $k$ occurrences of the subpath $UUU$.

Consider the map $\psi$ defined recursively, which transforms Dyck paths as follows:

$$\begin{aligned}
\psi(\epsilon) &= \epsilon,\\
\psi(UDP) &= UD\psi(P),\\
\psi(UUQ_1D\cdots UQ_kDDP) &= U^{k+1}D\psi(Q_1)D\psi(Q_2)D\cdots\psi(Q_k)D\psi(P),
\end{aligned}$$

where $k \geq 1$, and $P$, $Q_1, \ldots, Q_k$ are possibly empty Dyck paths.

We can easily verify by induction that $\psi$ is a bijection on Dyck paths. Moreover, $\psi$ maps the number of valleys of odd height to the number of occurrences of $UUU$. This claim can be proved by induction on the semilength. Let $\mathtt{dr}(R)$ denote the number of occurrences of $UUU$ in $R$. The statement is clear for the first two cases in the definition of $\psi$. For the third case, we have

$$\begin{aligned}\mathtt{dr}(\psi(UUQ_1D\cdots UQ_kDDP) &= \mathtt{dr}(U^{k+1}D\psi(Q_1)D\psi(Q_2)D\cdots\psi(Q_k)D\psi(P))\\ &= k-1+\mathtt{dr}(\psi(Q_1))+\cdots+\mathtt{dr}(\psi(Q_k)+\mathtt{dr}\psi(P)),\end{aligned}$$

which, by the induction hypothesis, equals $k-1$ plus the number of valleys of odd height in $Q_1, Q_2, \ldots, Q_k, P$. This quantity precisely matches the number of valleys of odd height in $UUQ_1D\cdots UQ_kDDP$, completing the induction.

For instance, if $R = U\ UD\ UUDUDD\ UD\ D\ UUDD$, then we have

$$\psi(R) = UUUUDDUDUDDDUUDD,$$

and the number of valleys of odd heights in $R$ is two, which is also the number of occurrences of $UUU$ in $\psi(R)$; see Figure 6.

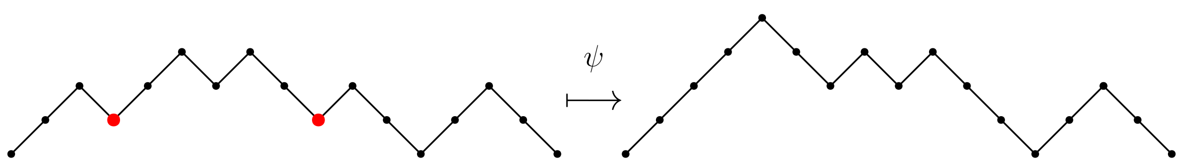



FIGURE 6. Example of the bijection $\psi$.

**Corollary 4.5.** *An asymptotic estimate for the expected number of valleys of even height (resp. odd height) in all Dyck paths of semilength $n$ is given by $n/4$ (resp. $n/4$).*

4.0.3. *The statistics of valleys at heights in arithmetical progressions.* In this section, we generalize the above result to sets defined by an arithmetic progression. Let $a$ and $b$ be integers, with $0 \leq b < a$. We define $A_{a,b} = \{m \in \mathbb{N} : m \equiv b \pmod{a}\}$ as the set of residue classes modulo $a$.

The following system of equations is obtained from Theorem 4.1.

$$\begin{cases} \nabla_{A_{a,0}}(x,y) = 1 + x\nabla_{A_{a,a-1}}(x,y) + xy\nabla_{A_{a,a-1}}(x,y)\nabla_{A_{a,0}}(x,y), \\ \nabla_{A_{a,1}}(x,y) = 1 + x\nabla_{A_{a,0}}(x,y)\nabla_{A_{a,1}}(x,y), \\ \qquad\vdots \\ \nabla_{A_{a,a-1}}(x,y) = 1 + x\nabla_{A_{a,a-2}}(x,y)\nabla_{A_{a,a-1}}(x,y). \end{cases}$$

The proof of the corollary follows from solving this system of equations for $a = 3$.

**Corollary 4.6.** *The number of Dyck paths with peaks at $A_{3,0}$, $A_{3,1}$, and $A_{3,2}$ is counted by generating functions satisfying the following equations:*

$$\begin{aligned}\nabla_{A_{3,0}}(x,y) &= \frac{(x-1)\,W + xy - x}{xy-1},\\ \nabla_{A_{3,1}}(x,y) &= W,\\ \nabla_{A_{3,2}}(x,y) &= \frac{(x-1)\,W + xy - x - y + 1}{2xy - x - y},\end{aligned}$$

*where $W$ is equal to*

$$\frac{-x^2y + x^2 + xy - 1 + \sqrt{x^4y^2 - 2x^4y - 2x^3y^2 + x^4 - 2x^3y + x^2y^2 + 6x^2y + 2x^2 - 2xy - 4x + 1}}{2x\,(x-1)}.$$

From Corollary 4.6, we obtain the first few rows of the corresponding Narayana arrays.

$$\mathcal{V}^{(1)}_{A_{3,0}} = \begin{pmatrix} 1 & 0 & 0 & 0 & 0 & 0 & 0\\ 1 & 0 & 0 & 0 & 0 & 0 & 0\\ 1 & 1 & 0 & 0 & 0 & 0 & 0\\ 2 & 2 & 1 & 0 & 0 & 0 & 0\\ 5 & 5 & 3 & 1 & 0 & 0 & 0\\ 13 & 15 & 9 & 4 & 1 & 0 & 0\\ 35 & 46 & 31 & 14 & 5 & 1 & 0\\ 97 & 143 & 108 & 54 & 20 & 6 & 1 \end{pmatrix}, \quad \mathcal{V}^{(1)}_{A_{3,1}} = \begin{pmatrix} 1 & 0 & 0 & 0 & 0 & 0\\ 1 & 0 & 0 & 0 & 0 & 0\\ 2 & 0 & 0 & 0 & 0 & 0\\ 4 & 1 & 0 & 0 & 0 & 0\\ 9 & 4 & 1 & 0 & 0 & 0\\ 22 & 14 & 5 & 1 & 0 & 0\\ 57 & 47 & 21 & 6 & 1 & 0\\ 154 & 155 & 83 & 29 & 7 & 1 \end{pmatrix},$$

$$\mathcal{V}^{(1)}_{A_{3,2}} = \begin{pmatrix} 1 & 0 & 0 & 0 & 0 & 0\\ 1 & 0 & 0 & 0 & 0 & 0\\ 2 & 0 & 0 & 0 & 0 & 0\\ 5 & 0 & 0 & 0 & 0 & 0\\ 13 & 1 & 0 & 0 & 0 & 0\\ 35 & 6 & 1 & 0 & 0 & 0\\ 97 & 27 & 7 & 1 & 0 & 0\\ 275 & 109 & 36 & 8 & 1 & 0\\ 794 & 415 & 165 & 46 & 9 & 1 \end{pmatrix}.$$

Notice that we have $\mathcal{V}^{(1)}_{A_{3,0}} = \mathcal{P}^{(1)}_{A_{3,2}}$ and $\mathcal{V}^{(1)}_{A_{3,1}} = \mathcal{P}^{(1)}_{A_{3,0}}$. We now define a bijection between Dyck paths with $k$ peaks of height congruent to 2 mod 3 (respectively, 0 mod 3) and Dyck paths with $k$ valleys of height congruent to 0 mod 3 (respectively, 1 mod 3). Given a Dyck path $R$ from the first set, we construct its image by reading $R$ from left to right and performing the following local transformations: if a peak $UD$ occurs at a height congruent to 2 mod 3 (respectively, 0 mod 3), it is replaced with a valley $DU$; if a valley $DU$ occurs at a height congruent to 0 mod 3 (respectively, 1 mod 3), it is replaced with a peak $UD$. For instance, if $R = UUD\ UUDUDD\ UD\ D\ UUDD$, then its image is $R' = UDUUUDUDDDUUDDUD$. The number of peaks of height congruent to 2 mod 3 in $R$ is three, which equals the number of valleys of height congruent to 0 mod 3 in $R'$. See Figure 7.

The array $\mathcal{V}^{(1)}_{A_{3,2}}$ does not appear in the OEIS. However, its first column is a shift of the first column of $\mathcal{V}^{(1)}_{A_{3,0}}$. The bijection is given as follows: a Dyck path of semilength $n$ with

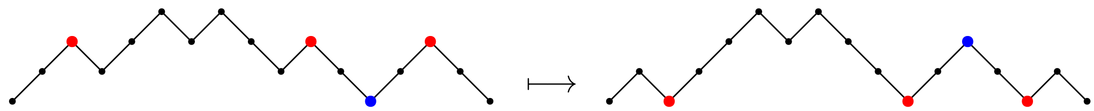


FIGURE 7. Example of the bijection.

no valleys at heights congruent to 0 mod 3 is of the form $UPD$, where $P$ is a Dyck path with no valleys at heights congruent to 2 mod 3.

**Corollary 4.7.** *An asymptotic estimate for the expected number of valleys of heights congruent to $i$ modulo 3 $(0 \leq i \leq 2)$ among all Dyck paths of semilength $n$ is given by $n/6$ (for each $i = 0, 1, 2$).*

## 5. Generalized Narayana Arrays of Type 2: The Peak Statistic

In Section 3, we enumerated Dyck paths with peaks at heights belonging to a given set $S$. In this section, we refine the count by considering Dyck paths with exactly $k$ peaks at heights in $S$. We begin by analyzing the special case $S = \{m\}$, providing bijections with compositions whose parts are of size at most $\ell$. We then study the case $S = \{m, m+1\}$, expressing the corresponding generating function as a finite continued fraction. In this case, we observe that the row sums of the associated Narayana arrays are related to a generalization of the Pell numbers.

For other choices of $S$, such as the set of positive even integers or the set of positive odd integers, we establish connections with Motzkin paths. We also examine cases where $S$ is an arithmetic progression, and we provide bijections with Dyck paths whose peaks occur at heights congruent to $i \pmod 3$, for $i = 0, 1, 2$.

### 5.1. The statistics of peaks at heights in a given set.

In this section, we count the number of Dyck paths with peaks at heights in set $S$.

Recall that $\overline{P}_S(n, k)$ denotes the number of Dyck paths of semilength $n$ with exactly $k$ peaks, where the height of each peak is contained in $S$. We introduce the bivariate generating function

$$\wedge_S(x, y) = \sum_{n,k \geq 0} \overline{P}_S(n, k) x^n y^k.$$

From the first return decomposition we obtain the following theorem.

**Theorem 5.1.** *The number of Dyck paths with exactly $k$ peaks whose heights lie in a subset $S \subseteq \mathbb{Z}^+$ is counted by the generating function $\wedge_S(x, y)$, which satisfies the following functional equation:*

$$\wedge_S(x, y) = \begin{cases} 1 + x(\wedge_{S_-}(x, y) - 1)\wedge_S(x, y), & \text{if } 1 \notin S; \\ 1 + x(\wedge_{S_-}(x, y) - 1)\wedge_S(x, y) + xy\wedge_S(x, y), & \text{otherwise.} \end{cases}$$

*Moreover, $\wedge_\emptyset(x, y) = 1$.*

*Proof.* Let $R$ be a non-empty Dyck path, and write its first return decomposition as $R = UPDQ$, where $P$ and $Q$ are (possibly empty) Dyck paths.

Suppose $1 \in S$. We distinguish two cases:

**Case** 1. If $R = UDQ$, then $R \setminus Q = UD$ has exactly one peak at height $1 \in S$, and $Q$ contributes the remaining structure. Thus, the contribution of $R$ to the generating function is $1 + xy \wedge_S(x, y)$.

**Case** 2. If $R = UPDQ$ with $P$ non-empty, then the number of peaks of $R$ at heights in $S$ is $a+b$, where $a$ (respectively, $b$) denotes the number of peaks in $P$ (respectively, $Q$) at heights in $S_-$ (respectively, $S$), with $S_- = \{h-1 : h \in S\}$. Thus, the contribution of $R$ to the generating function is $x(\wedge_{S_-}(x, y) - 1)\wedge_S(x, y)$.

Combining both cases, we obtain the following functional equation:

$$\wedge_S(x, y) = 1 + xy\wedge_S(x, y) + x(\wedge_{S_-}(x, y) - 1)\wedge_S(x, y).$$

If instead $1 \notin S$, then the decomposition $R = UPDQ$ with $P$ non-empty leads to the functional equation:

$$\wedge_S(x, y) = 1 + x(\wedge_{S_-}(x, y) - 1)\Delta_S(x, y).$$

Finally, observe that if $S = \emptyset$, then $\wedge_\emptyset = 1$, since the empty path is the unique Dyck path with no peaks. □

5.1.1. *The statistics of peaks at heights in $S = \{m\}$.* The arrays $\mathcal{P}^{(2)}_{\{2\}}$ and $\mathcal{P}^{(2)}_{\{3\}}$ correspond to the sequences A030528 and A078803, respectively. More generally, the Narayana array of type 2, denoted $\mathcal{P}^{(2)}_{\{\ell\}}$, $\ell \geq 1$, is associated with integer compositions. A *composition* of a positive integer $n$ is a sequence of positive integers, called *parts*, whose sum is $n$.

We define a map $\omega$ from the set of Dyck paths enumerated by $\overline{P}_{[\ell]}(n, k)$ to the set of integer compositions of $n - \ell$ into $k - 1$ parts, where each part is at most $\ell$. Specifically, any Dyck path $P$ counted by $\overline{P}_{[\ell]}(n, k)$ can be written as

$$P = U^{s_1}D^{r_1}U^{s_2}D^{r_2}\cdots U^{s_k}D^{r_k},$$

where $s_1 + s_2 + \cdots + s_k = n = r_1 + r_2 + \cdots + r_k$, $s_1 = \ell = r_k$, and $1 \leq s_i, r_i \leq \ell$ for all $i$. We define the function $\omega(P) = s_2 + \cdots + s_k$. It is clear that $\omega(P)$ yields a composition of $n - \ell$ into $k - 1$ parts, each contained in $[\ell]$, and it establishes a bijection.

For example, for the Dyck paths of semilength 6 with all peaks of height 3, we have the following correspondence:

$$\omega(UUUDDDUUUDDD) = 3, \qquad \omega(UUUDDUUDUDDD) = 2 + 1,$$
$$\omega(UUUDUDDUUDDD) = 1 + 2, \qquad \omega(UUUDUDUDUDDD) = 1 + 1 + 1.$$

5.1.2. *The statistics of peaks at heights in $S = \{m, m+1\}$.* In this section, we study the special case where the Dyck paths have exactly $k$ peaks, each of whose heights lies in $S = \{m, m+1\}$, which corresponds to a generating function represented by a finite continued fraction. When $S = \{2, 3\}$, the associated generating function yields the Narayana numbers of type 2.

The proof of the following result is a direct consequence of Theorem 5.1.

**Corollary 5.2.** *The number of Dyck paths with exactly k peaks, each of whose heights lies in $\{m, m+1\}$, is counted by the generating function $\wedge_{\{m,m+1\}}(x,y)$, given by the continued fraction*

$$\wedge_{\{m,m+1\}}(x,y) = \left.\cfrac{1}{1+x-\cfrac{x}{1+x-\cfrac{x}{\cfrac{\vdots}{1+x-\cfrac{x}{1-xy+x-\cfrac{x^2y}{1-xy}}}}}}\right\} m \text{ times},$$

*which implies that*

$$\wedge_{\{m,m+1\}}(x,1) = 1 + \frac{x^m}{1-2x-x^2-x^3-\cdots-x^m}.$$

For instance, if we set $m=2$, that is $S=\{2,3\}$, then Corollary 5.2 provides

$$\wedge_{\{2,3\}}(x,y) = \frac{x^2y^2 - x^2y - 2xy + 1}{x^3y^2 - x^3y + x^2y^2 - 2x^2y - 2xy + 1}.$$

From this generating function we can obtain the first few rows of the Narayana array of type 2 associated with the set $S=\{2,3\}$.

$$\mathcal{P}^{(2)}_{\{2,3\}} = \begin{pmatrix} 1 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 & 0 & 0 \\ 0 & 1 & 1 & 0 & 0 & 0 & 0 \\ 0 & 0 & \boxed{\mathbf{4}} & 1 & 0 & 0 & 0 \\ 0 & 0 & 3 & 8 & 1 & 0 & 0 \\ 0 & 0 & 1 & 14 & 13 & 1 & 0 \\ 0 & 0 & 0 & 12 & 38 & 19 & 1 \end{pmatrix}.$$

For example, $\overline{P}_{\{2,3\}}(4,2) = 4$, which corresponds to the boxed entry in the array $\mathcal{P}^{(2)}_{\{2,3\}}$ above. The 4 Dyck paths corresponding to this value are illustrated in Figure 1. This array does not appear in the OEIS, but its row sums correspond to the Pell sequence A000129.

More generally, we can prove that the row sums of the Narayana array $\mathcal{P}^{(2)}_{\{m,m+1\}}$ correspond to the $m$-generalized Pell numbers studied by Bravo et al. in [4]. Moreover, these sequences also count generalized bicolored compositions of $n-m$ with parts in the set $[m]$; that is, compositions of $n$ into parts from the set $\{1, 2, \dots, m\}$ in which the part 1 may appear in two distinct colors. These two colors of the part 1 are denoted by the subscripts $1_1$ and $1_2$.

**Open Problem 5.1.** It would be interesting to identify a statistic on bicolored compositions that corresponds to the number of peaks in Dyck paths whose peaks all lie at heights in $\{m, m+1\}$.

**Corollary 5.3.** *An asymptotic estimate for the expected number of peaks at heights in* $\{2, 3\}$ *among all Dyck paths of semilength* $n$ *having peaks at heights* 2 *or* 3 *is*

$$\frac{4-\sqrt{2}}{4} \cdot n.$$

5.1.3. *The statistics of peaks at either even or odd heights.* We now use again the notation introduced in Section 3.1.3. Specifically, we let $\mathbb{E}$ denote the set of all non-negative even integers, and $\mathbb{O}$ denote the set of all positive odd integers.

**Corollary 5.4.** *We have*

$$\begin{aligned}\wedge_{\mathbb{E}}(x,y) =& \frac{1+x(2-y)+x^2(1-y)-\sqrt{(1+x)(1+x-xy)(1-2x+x^2-xy-x^2y)}}{2x(1+x)},\\ \wedge_{\mathbb{O}}(x,y) =& \frac{1+x}{1+x(1-y)}\wedge_{\mathbb{E}}(x,y).\end{aligned}$$

The proof of this corollary follows by solving the system of equations derived from Theorem 5.1.

$$\begin{cases}\wedge_{\mathbb{E}}(x,y) = & 1+x(\wedge_{\mathbb{O}}(x,y)-1)\wedge_{\mathbb{E}}(x,y),\\ \wedge_{\mathbb{O}}(x,y) = & 1+x(\wedge_{\mathbb{E}}(x,y)-1)\wedge_{\mathbb{O}}(x,y)+xy\wedge_{\mathbb{O}}(x,y).\end{cases}$$

From Corollary 5.4, we obtain the first few rows of the Narayana arrays of type 2 associated with the two sets of even and odd integers.

$$\mathcal{P}_{\mathbb{E}}^{(2)} = \begin{pmatrix}1&0&0&0&0&0&0\\0&0&0&0&0&0&0\\0&1&0&0&0&0&0\\0&0&1&0&0&0&0\\0&1&1&1&0&0&0\\0&0&3&2&1&0&0\\0&1&3&7&3&1&0\\0&0&6&12&13&4&1\end{pmatrix} \quad\text{and}\quad \mathcal{P}_{\mathbb{O}}^{(2)} = \begin{pmatrix}1&0&0&0&0&0&0&0\\0&1&0&0&0&0&0&0\\0&0&1&0&0&0&0&0\\0&1&0&1&0&0&0&0\\0&0&3&0&1&0&0&0\\0&1&1&6&0&1&0&0\\0&0&6&4&10&0&1&0\\0&1&3&21&10&15&0&1\end{pmatrix}.$$

We observe that the row sums of $\mathcal{P}_{\mathbb{E}}^{(2)}$ correspond to the Riordan numbers A005043. The Riordan numbers have several combinatorial interpretations; see, for example [6]. One such interpretation is the number of Motzkin paths that avoid horizontal steps at ground level. Recall that a *Motzkin path* is a path in the first quadrant of the $xy$-plane having up-steps ($U=(1,1)$), down-steps ($D=(1,-1)$), and horizontal steps $H=(1,0)$. These paths start at the origin, end on the $x$-axis, and never pass below it. We prove that the array $\mathcal{P}_{\mathbb{E}}^{(2)}$ provides a refinement of the Riordan numbers.

We define $\mathcal{L}(n,k)$ to be the set of all Motzkin paths of length $n$ that avoid horizontal steps at a ground level and contain exactly $k-1$ valleys of the form $DU$ along with horizontal steps. For example, the Motzkin path $P = UUHDUDHD$ has one valley and two horizontal steps. Therefore, $P \in \mathcal{L}(8,3)$.

Let $\mathcal{W}_{\mathbb{E}}(n,k)$ denote the set of Dyck path of semilength $n$ with exactly $k$ peaks, all of which have even height. For the bijection that follows, we use the construction provided by Callan in [18, A001006]. Since $P \in \mathcal{W}_{\mathbb{E}}(n,k)$ can be viewed as a word of length $2n$, it can be divided into consecutive subwords of length 2.

Define the function $\phi : \mathcal{W}_{\mathbb{E}}(n,k) \to \mathcal{L}(n,k)$, which maps a Dyck path $P \in \mathcal{W}_{\mathbb{E}}(n,k)$, partitioned as described above, to a path $Q \in \mathcal{L}(n,k)$ by assigning $UU \to U$, $DD \to D$, and $DU \to H$.

For a path $P \in \mathcal{W}_{\mathbb{E}}(n,k)$, we note the following facts: any subpath $D^rU^t$ of $P$ between peak vertices has an even length. Therefore, $r$ and $t$ must have the same parity. Under $\phi$, such a subpath is mapped to $D^kH^iU^j$, where $i=1$ if $r$ is odd, and $i=0$ if $r$ is even, with $k, j \geq 0$. Consequently, every valley in $P$ is mapped under $\phi$ to either $H$ or $DU$. Moreover, the absence of peaks with odd heights in $P$, implies that $\phi$ does not map subpaths of $P$ to horizontal steps at ground level in $Q$. For example,

$$\phi(UUUUDDDUUUDDUUDDDDUUDD) = UUDHUDUDDUD,$$

see Figure 8.

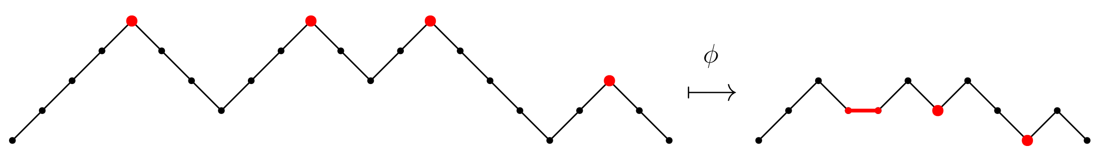



FIGURE 8. Example of the bijection $\phi$.

Moreover, the array $\mathcal{P}_{\mathbb{O}}^{(2)}$ is a refinement of the Motzkin numbers. There is a bijection between Dyck paths of semilength $n$ with $k$ peaks, all of which have odd height, and Motzkin paths having exactly $k$ occurrences of either $H$ or $DU$.

Given such a Dyck path, we construct the corresponding Motzkin path as follows: first, delete the initial $U$ and the final $D$ from the Dyck path. Then, partition the remaining Dyck word into consecutive subwords of length 2, and apply the transformation $UU \mapsto U$, $DD \mapsto D$, and $DU \mapsto H$.

**Corollary 5.5.** *An asymptotic estimate for the expected number of peaks at even heights (respectively, at odd heights) among all Dyck paths of semilength $n$ whose peaks are at even heights (respectively, at odd) is given by $4n/9$ in both cases.*

5.1.4. *The statistics of peaks at heights in arithmetic progression.* We now recall the notation introduced in Section 3.1.3. Specifically, let $a$ and $b$ be integers with $0 \leq b < a$, and define $A_{a,b} = \{m \in \mathbb{N} : m \equiv b \pmod{a}\}$ as the set of non-negative integers congruent to $b$ modulo $a$.

From Theorem 5.1, we obtain the following system of equations:

$$\begin{cases} \wedge_{A_{a,0}}(x,y) = 1 + x(\wedge_{A_{a,a-1}}(x,y) - 1)\wedge_{A_{a,0}}(x,y), \\ \wedge_{A_{a,1}}(x,y) = 1 + x(\wedge_{A_{a,0}}(x,y) - 1)\wedge_{A_{a,1}}(x,y) + xy\wedge_{A_{a,1}}(x,y), \\ \wedge_{A_{a,2}}(x,y) = 1 + x(\wedge_{A_{a,1}}(x,y) - 1)\wedge_{A_{a,2}}(x,y), \\ \vdots \\ \wedge_{A_{a,a-1}}(x,y) = 1 + x(\wedge_{A_{a,a-2}}(x,y) - 1)\wedge_{A_{a,a-1}}(x,y). \end{cases}$$

The proof of the following corollary is obtained by solving the above system of equations for the case $a = 3$.

**Corollary 5.6.** *We have*

$$\wedge_{A_{3,0}}(x,y) = \frac{1}{1 + x - xR(x,y)},$$
$$\wedge_{A_{3,1}}(x,y) = \frac{-1 + (x+1)R(x,y)}{xR(x,y)},$$
$$\wedge_{A_{3,2}}(x,y) = R(x,y).$$

*with* $R(x,y)$ *equal to*

$$\frac{x^3y - x^3 + 3\,x^2y - 2\,x^2 + xy - 2\,x - 1 + \sqrt{(x^2+x+1)\,(x^4y^2 - 2\,x^4y + x^3y^2 + x^4 + x^2y^2 - x^3 - 2\,xy - x + 1)}}{2x\,(x^2y - x^2 + xy - x - 1)}.$$

From the generating functions in Corollary 5.6, we obtain the first few rows of the corresponding Narayana arrays.

$$\mathcal{P}^{(2)}_{A_{3,0}} = \begin{pmatrix} 1&0&0&0&0&0&0 \\ 0&0&0&0&0&0&0 \\ 0&0&0&0&0&0&0 \\ 0&1&0&0&0&0&0 \\ 0&0&1&0&0&0&0 \\ 0&0&1&1&0&0&0 \\ 0&1&1&2&1&0&0 \\ 0&0&3&3&3&1&0 \\ 0&0&3&8&6&4&1 \end{pmatrix}, \quad \mathcal{P}^{(2)}_{A_{3,1}} = \begin{pmatrix} 1&0&0&0&0&0&0&0 \\ 0&1&0&0&0&0&0&0 \\ 0&0&1&0&0&0&0&0 \\ 0&0&0&1&0&0&0&0 \\ 0&1&0&0&1&0&0&0 \\ 0&0&3&0&0&1&0&0 \\ 0&0&1&6&0&0&1&0 \\ 0&1&1&4&10&0&0&1 \end{pmatrix},$$

$$\mathcal{P}^{(2)}_{A_{3,2}} = \begin{pmatrix} 1&0&0&0&0&0&0 \\ 0&0&0&0&0&0&0 \\ 0&1&0&0&0&0&0 \\ 0&0&1&0&0&0&0 \\ 0&0&1&1&0&0&0 \\ 0&1&0&2&1&0&0 \\ 0&0&3&1&3&1&0 \\ 0&0&3&6&3&4&1 \end{pmatrix}.$$

We observe that the row sums of $\mathcal{P}^{(2)}_{A_{3,0}}$ (respectively, $\mathcal{P}^{(2)}_{A_{3,1}}$, $\mathcal{P}^{(2)}_{A_{3,2}}$) generate the sequence A166300 (respectively, A004148, A004148).

Let us define recursively three bijections $f_0$, $f_1$, and $f_2$ from the sets $A_n$, $B_{n-1}$, and $C_{n-1}$, respectively, where:

- $A_n$ denotes the set of Dyck paths of semilength $n$ with all peaks at height congruent to 0 mod 3,
- $B_{n-1}$ denotes the set of Dyck paths of semilength $n-1$ with all peaks at height congruent to 1 mod 3,
- $C_{n-1}$ denotes the set of Dyck paths of semilength $n-1$ with all peaks at height congruent to 2 mod 3.

These bijections map to peakless Motzkin paths of length $n$ (for $f_0$) and of length $n-1$ (for $f_1$ and $f_2$).

$$\begin{aligned}
f_0(UPDQ) &= Uf_2(P)Df_0(Q) && \\
f_1(UD) &= \epsilon && \\
f_1(UDQ) &= Hf_1(Q) && \text{if } Q \neq \epsilon \\
f_1(UPD) &= f_0(P) && \text{if } P \neq \epsilon \\
f_1(UPDQ) &= f_0(P)Hf_1(Q) && \text{if } P, Q \neq \epsilon \\
f_2(UPD) &= Hf_1(P) && \\
f_2(UPDQ) &= Uf_1(Q)Df_2(P) && \text{if } Q \neq \epsilon.
\end{aligned}$$

For instance, if $R = UUUDUDDUUDDUUUDDD \in A_9$, then we have

$$\begin{aligned}
f_0(R) &= Uf_2(UUDUDDUUDD)Df_0(UUUDDD) \\
&= UUf_2(UUDD)Df_1(UDUD)DUHD = UUHDHDUHD.
\end{aligned}$$

**Open Problem 5.2.** We were unable to find bijections that map the number of peaks to a specific statistic on peakless Motzkin paths. Therefore, we leave it as an open question.

**Corollary 5.7.** *An asymptotic estimate for the expected number of peaks of heights congruent to $i \bmod 3$ $(0 \leq i \leq 2)$, in all Dyck paths of semilength $n$ with peaks at heights $i \bmod 3$ is given by, for each $i = 0, 1, 2$, the following expression:*

$$4n \cdot \left(1 - \frac{2}{\sqrt{5}}\right).$$

## 6. Generalized Narayana Arrays of Type 2: The Valley Statistic

In Section 4, we enumerated Dyck paths with a given number of valleys at heights belonging to a given set $S$. In this section, we refine the count by considering Dyck paths with exactly $k$ valleys, and where every valley has height in $S$. We begin with the case $S = \{m, m+1\}$, and we observe that the associated Narayana array corresponds to a shifted version of the Pascal triangle. For other choices of $S$, such as the set of positive even integers or the set of positive odd integers, we establish connections with Motzkin paths containing a fixed number of peaks and horizontal steps. As in previous sections, we also examine the case where $S$ is an arithmetic progression, and we note that the row sum of the corresponding Narayana array is related to the set of peakless Motzkin paths.

Finally, for the case $S = [\ell]$, we find an interesting bijection with a full binary tree that avoids a specific subtree in which every right child is also a left child.

We introduce the generating function

$$\vee_S(x,y) = \sum_{n,k\geq 0} \overline{V}_S(n,k)x^n y^k,$$

where $\overline{V}_S(n,k)$ denotes the number of Dyck paths of semilength $n$ with exactly $k$ valleys.

Applying the same reasoning as in Theorem 5.1, we obtain the bivariate generating function for the dual Narayana array of type 2.

**Theorem 6.1.** *Let $S \subseteq \mathbb{N}$. The generating functions $\vee_S(x,y)$ satisfies*

$$\vee_S(x,y) = \begin{cases} 1 + x\vee_{S^*_-}(x,y), & \text{if } 0 \notin S; \\ 1 + x\vee_{S^*_-}(x,y) + xy\vee_{S^*_-}(x,y)(\vee_S(x,y) - 1), & \text{otherwise.} \end{cases}$$

*Moreover, $\vee_\emptyset(x,y) = \frac{1}{1-x}$.*

6.0.1. *The statistics of valleys at heights in $S = \{m\}$ and $S = \{m, m+1\}$.* Whenever $S = \{m\}$, with $m \geq 0$, we obtain

$$\vee_{\{m\}}(x,y) = x^m \vee_{\{0\}}(x,y) = \frac{(xy-1)x^m}{xy + x - 1}.$$

Additionally, the array $\mathcal{V}^{(2)}_{\{m\}}$ is a shift of the well-known Pascal triangle.

Whenever $S = \{m, m+1\}$, with $m \geq 0$, we obtain

$$\vee_{\{m,m+1\}}(x,y) = x^m \vee_{\{0,1\}}(x,y) = x^m(1 + x\vee_{\{0\}}(x,y)) = x^m \frac{x^2y + xy - 1}{xy + x - 1}.$$

Additionally, the array $\mathcal{V}^{(2)}_{\{m,m+1\}}$ is a shift of the well-known Pascal triangle.

An asymptotic estimate for the expected number of valleys at heights in $m$ (respectively, $\{m, m+1\}$), in all Dyck paths of semilength $n$ having all valleys at height $m$ (respectively, $\{m, m+1\}$) is given by $n/2$.

6.0.2. *The statistics of valleys at either even or odd heights.* In this part, we examine the two cases $S = \mathbb{E}$ and $S = \mathbb{O}$, corresponding to the sets of even and odd non-negative integers, respectively. From Theorem 6.1 we derive the following generating functions:

**Corollary 6.2.** *We have*

$$\vee_{\mathbb{E}}(x,y) = \frac{1 + x(2-y) - \sqrt{1 - 2xy - x^2(4-y^2)}}{2x}, \qquad \vee_{\mathbb{O}}(x,y) = 1 + x\vee_{\mathbb{E}}(x,y).$$

From the generating functions we obtain the first few rows of the Narayana arrays of type 2 associated with the set of even and odd positive integers:

$$\mathcal{V}_{\mathbb{E}}^{(2)} = \begin{pmatrix} 1 & 0 & 0 & 0 & 0 & 0 & 0 \\ 1 & 0 & 0 & 0 & 0 & 0 & 0 \\ 1 & 1 & 0 & 0 & 0 & 0 & 0 \\ 1 & 2 & 1 & 0 & 0 & 0 & 0 \\ 1 & 4 & 3 & 1 & 0 & 0 & 0 \\ 1 & 6 & 9 & 4 & 1 & 0 & 0 \\ 1 & 9 & 19 & 16 & 5 & 1 & 0 \\ 1 & 12 & 38 & 44 & 25 & 6 & 1 \end{pmatrix}, \qquad \mathcal{V}_{\mathbb{O}}^{(2)} = \begin{pmatrix} 1 & 0 & 0 & 0 & 0 & 0 & 0 \\ 1 & 0 & 0 & 0 & 0 & 0 & 0 \\ 1 & 0 & 0 & 0 & 0 & 0 & 0 \\ 1 & 1 & 0 & 0 & 0 & 0 & 0 \\ 1 & 2 & 1 & 0 & 0 & 0 & 0 \\ 1 & 4 & 3 & 1 & 0 & 0 & 0 \\ 1 & 6 & 9 & 4 & 1 & 0 & 0 \\ 1 & 9 & 19 & 16 & 5 & 1 & 0 \end{pmatrix}.$$

Note that the row sum of $\mathcal{V}_{\mathbb{E}}^{(2)}$ is equal to the Motzkin numbers A001006, and $\mathcal{V}_{\mathbb{O}}^{(2)}$ is a shift of $\mathcal{V}_{\mathbb{E}}^{(2)}$. Therefore, the dual Narayana array of type two $\mathcal{V}_{\mathbb{E}}^{(2)}$ is a refinement of the Motkzin numbers. In fact, we define $\mathcal{M}(n,k)$ as the set of all Motzkin paths of length $n$ that contain exactly $k$ peaks $UD$ along with horizontal steps. Let $\mathcal{J}_{\mathbb{E}}(n,k)$ denote the set of Dyck paths of semilength $n$ with exactly $k$ valleys, all of which have even heights. Define the function $\overline{\phi} : \mathcal{J}_{\mathbb{E}}(n,k) \to \mathcal{M}(n,k+1)$, which maps a Dyck path $P \in \mathcal{J}_{\mathbb{E}}(n,k)$, partitioned into consecutive subwords of length 2, to a path $Q \in \mathcal{M}(n,k+1)$ by assigning $UU \mapsto U$, $DD \mapsto D$, and $UD \mapsto H$.

For a path $P \in \mathcal{J}_{\mathbb{E}}(n,k)$, we note the following facts: any subpath $U^rD^t$ of $P$ between valley vertices has an even length. Therefore, $r$ and $t$ must have the same parity. Under $\overline{\phi}$, such a subpath is mapped to $D^kH^iU^j$, where $i=1$ if $r$ is odd and $i=0$ if $r$ is even, with $k,j \geq 0$. Consequently, every valley in $P$ is mapped under $\overline{\phi}$ to either $H$ or $DU$.

For example,

$$\overline{\phi}(UUUUDDDDUUUDUDDDUUUDDD) = UUDDUHHDUHD,$$

see Figure 9.

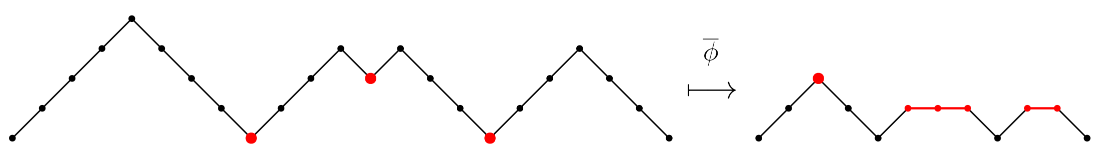


FIGURE 9. Example of the bijection $\overline{\phi}$.

It worth noting that $\mathcal{M}(n,k)$ also represents all Motzkin paths of length $n$ with exactly $k$ weak valleys (that is, with steps of the form: $DU$, $HU$, $DH$, $HH$). See the sequence A110470.

**Corollary 6.3.** *An asymptotic estimate for the expected number of valleys at even heights (respectively, odd heights) in all Dyck paths of semilength $n$ with all valleys at even heights (respectively, odd heights) is given by $n/3$ for both cases.*

6.0.3. *The statistics of valleys at heights in $S = \{m \in \mathbb{N} : m \equiv b \pmod 3\}$.* In this section, we generalize the previous results to sets defined by arithmetic progressions. As before, we provide generating functions that enumerate Dyck paths with exactly $k$ valleys, each of whose heights lies within these sets.

We now recall the notation introduced before. Specifically, let $a$ and $b$ be integers with $0 \leq b < a$, and define $A_{a,b} = \{m \in \mathbb{N} : m \equiv b \pmod{a}\}$ as the set of non-negative integers congruent to $b$ modulo $a$. From Theorem 6.1, we obtain the following system of equations:

$$\begin{cases} \vee_{A_{a,0}}(x,y) = 1 + x\vee_{A_{a,a-1}}(x,y) + xy(\vee_{A_{a,a-1}}(x,y) - 1)\vee_{A_{a,0}}(x,y), \\ \vee_{A_{a,1}}(x,y) = 1 + x(\vee_{A_{a,0}}(x,y), \\ \vee_{A_{a,2}}(x,y) = 1 + x\vee_{A_{a,1}}(x,y), \\ \qquad\vdots \\ \vee_{A_{a,a-1}}(x,y) = 1 + x\vee_{A_{a,a-2}}(x,y). \end{cases}$$

The proof of the following corollary is obtained by solving the above system of equations for the case $a = 3$.

**Corollary 6.4.** *We have*

$$\begin{aligned} \vee_{A_{3,0}}(x,y) &= R(x,y), \\ \vee_{A_{3,1}}(x,y) &= 1 + xR(x,y), \\ \vee_{A_{3,2}}(x,y) &= 1 + x + x^2R(x,y), \end{aligned}$$

*with* $R(x,y)$ *equal to*

$$\frac{1 + x^3y - x^3 - x^2y - xy - \sqrt{(x^2+x+1)\,(x^4y^2 - 2\,x^4y + x^3y^2 + x^4 + x^2y^2 - x^3 - 2\,xy - x + 1)}}{2x^3y}.$$

From these generating functions in Corollary 6.4, we obtain the first few rows of the corresponding Narayana arrays.

$$\mathcal{V}^{(2)}_{A_{3,0}} = \begin{pmatrix} 1&0&0&0&0&0&0&0 \\ 1&0&0&0&0&0&0&0 \\ 1&1&0&0&0&0&0&0 \\ 1&2&1&0&0&0&0&0 \\ 1&3&3&1&0&0&0&0 \\ 1&5&6&4&1&0&0&0 \\ 1&7&13&10&5&1&0&0 \\ 1&9&24&26&15&6&1&0 \\ 1&12&39&59&45&21&7&1 \end{pmatrix}, \quad \mathcal{V}^{(2)}_{A_{3,1}} = \begin{pmatrix} 1&0&0&0&0&0&0&0 \\ 1&0&0&0&0&0&0&0 \\ 1&0&0&0&0&0&0&0 \\ 1&1&0&0&0&0&0&0 \\ 1&2&1&0&0&0&0&0 \\ 1&3&3&1&0&0&0&0 \\ 1&5&6&4&1&0&0&0 \\ 1&7&13&10&5&1&0&0 \\ 1&9&24&26&15&6&1&0 \end{pmatrix},$$

$$\mathcal{V}^{(2)}_{A_{3,2}} = \begin{pmatrix} 1&0&0&0&0&0&0&0 \\ 1&0&0&0&0&0&0&0 \\ 1&0&0&0&0&0&0&0 \\ 1&0&0&0&0&0&0&0 \\ 1&1&0&0&0&0&0&0 \\ 1&2&1&0&0&0&0&0 \\ 1&3&3&1&0&0&0&0 \\ 1&5&6&4&1&0&0&0 \\ 1&7&13&10&5&1&0&0 \end{pmatrix}.$$

The two arrays $\mathcal{V}^{(2)}_{A_{3,1}}$ and $\mathcal{V}^{(2)}_{A_{3,2}}$ are quite similar to $\mathcal{V}^{(2)}_{A_{3,0}}$ up to a shift. The row sum of the array $\mathcal{V}^{(2)}_{A_{3,0}}$ corresponds to the sequence A004148, where the general term counts the number of peakless Motzkin numbers, where the general term counts the number of peakless Motzkin paths.

Below, we present a bijection between the set of Dyck paths of semilength $n$ with all valleys at heights congruent to 0 modulo 3, and the set of peakless Motzkin paths of length $n+1$:

$$f(\epsilon) = H$$
$$f(UDQ) = Hf(Q)$$
$$f(UUUPDDDQ) = Uf(P)Df(Q).$$

By a simple observation, the bijection maps the number of valleys to the number of occurrences of $DU$, $DH$, $DU$ and $HH$.

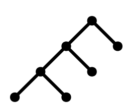
**Corollary 6.5.** *An asymptotic estimate for the expected number of valleys at heights* $i \bmod 3$ *(*$0 \le i \le 2$*) in all Dyck paths of semilength* $n$ *is given by (for each* $i = 0, 1, 2$*)*

$$0.2912966357 \cdot n.$$

6.0.4. *The statistics of valleys having at height in* $S = \{1, 2, 3, \ldots, \ell\}$. The row sums of the arrays $\mathcal{V}^{(2)}_{\{1,2,3\}}$, $\mathcal{V}^{(2)}_{\{1,2,3,4\}}$, $\mathcal{V}^{(2)}_{\{1,2,3,4,5\}}$ corresponds to the sequences A007051, A080937, A024175, respectively. More generally, we define a recursive bijection $\nu$ from Dyck paths with $k$ valleys of heights in $\{1, 2, \ldots, \ell\}$, for $\ell \ge 1$, to full binary trees that avoid a specific non-consecutive pattern $c_\ell$.

A *full binary tree* is a rooted, ordered tree in which each node has either zero or two children. Let $\mathcal{T}$ denote the set of all nonempty full binary trees. Given $P \in \mathcal{T}$, we say that a full binary tree $T$ *contains* $P$ as a (non-contiguous) tree pattern if $P$ can be obtained from $T$ through a finite sequence of edge contractions. Conversely, we say that $T$ *avoids* $P$ if it does not contain $P$ in this sense. For further details on the notion of pattern avoidance, we refer the reader to [5].

We denote by $c_\ell$ the unique full binary tree with $\ell + 1 \ge 1$ leaves such that every right child is a leaf. For example, $c_3$ is depicted in the following figure:



Let $\mathcal{T}(c_\ell)$ denote the set of trees in $\mathcal{T}$ that avoid the pattern $c_\ell$. Note that any tree $T \in \mathcal{T}(c_\ell)$ is necessarily either a single leaf $T = \bullet$, or a tree with a left child in $\mathcal{T}(c_\ell)$ and a right child in $\mathcal{T}(c_\ell)$. This recursive structure implies that the generating function for the number of trees with $n$ leaves in $\mathcal{T}(c_{\ell+1})$ satisfies the same functional equation as the generating function for the number of Dyck paths of semilength $n$ in $\mathcal{J}_{[\ell]} \backslash \{\epsilon\}$, where $\mathcal{J}_{[\ell]}$ denotes the set of Dyck paths in which all valleys occur at heights contained in $[\ell]$.

Let $P$ be a Dyck path in $\mathcal{J}_{[\ell]} \backslash \{\epsilon\}$. Then $P$ is of one of the following forms:

(*i*) $P = UD$, or

(*ii*) $P = UUP_1DUP_2D \cdots UP_rDD$ where $r \geq 1$ and $UP_iD$, $1 \leq i \leq r$, are Dyck paths in $\mathcal{J}_{[\ell-1]}$.

We define the map $\nu\colon \mathcal{J}_{[\ell]} \setminus \{\epsilon\} \to \mathcal{T}(c_{\ell+1})$ recursively as follows. Set $\nu(UD) = \bullet$, the full binary tree consisting of a single node. For a path of the form

$$P = UUP_1D\,UP_2D\,\cdots\,UP_rDD,$$

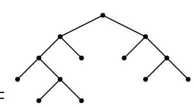
define $\nu(P)$ to be the full binary tree whose left child is $\nu(UP_1D)$ and whose right child is

$$\begin{cases} \nu(UUP_2D\,\cdots\,UP_rDD), & \text{if } r \geq 2, \\ \bullet, & \text{if } r = 1. \end{cases}$$

The map $\nu$ sends the semilength of a Dyck path to the number of leaves in the corresponding full binary tree. Moreover, $\nu$ maps the number of valleys of a given height $h \geq 1$ to the number of $h$-nodes whose right child is neither empty nor a leaf. Here, an *h-node* in a tree is a node such that the path from the root to that node contains exactly $h$ left children.

It follows that $\nu$ maps the set $\mathcal{J}_{[\ell]} \setminus \{\epsilon\}$ into $\mathcal{T}(c_{\ell+1})$. A cardinality argument (see [5]) implies that $\nu$ is a bijection.

For example, for a path in $\mathcal{J}_{[3]}\setminus\{\epsilon\}$ we have

$\nu(UUUUDUDDDUDUDDD) =$ 

Note that this tree avoids $c_4$. In Figure 10, we show how the mapping $\nu$ operates.

## 7. Appendix. A computational approach

This paper is accompanied by a *Mathematica* package that implements Theorem 3.1, providing a computational framework to derive the bivariate generating function $D_S(x, y)$. Our approach follows the methodology developed by Bu and Zeilberger in [8, 10].

We explain the methodology with the set $S = \{2, 3\}$. First, from Theorem 3.1 we obtain the following system of equations

$$\begin{aligned} D_{\{2,3\}}(x, y) &= 1 + xD_{\{1,2\}}(x, y)D_{\{2,3\}}(x, y) \\ D_{\{1,2\}}(x, y) &= 1 + x(D_{\{1\}}(x, y) - 1)D_{\{1,2\}}(x, y) + xyD_{\{1,2\}}(x, y) \\ D_{\{1\}}(x, y) &= 1 + x(C(x) - 1)D_{\{1\}}(x, y) + xyD_{\{1\}}(x, y). \end{aligned} \tag{1}$$

From the first two equations, we can eliminate the variable $D_{\{1,2\}}(x, y)$. In this case, we obtain the equation:

$$1 - D_{\{2,3\}}(x, y) + x - D_{\{1\}}(x, y)x + D_{\{1\}}(x, y)D_{\{2,3\}}(x, y) - xy + D_{\{2,3\}}(x, y)xy = 0. \tag{2}$$

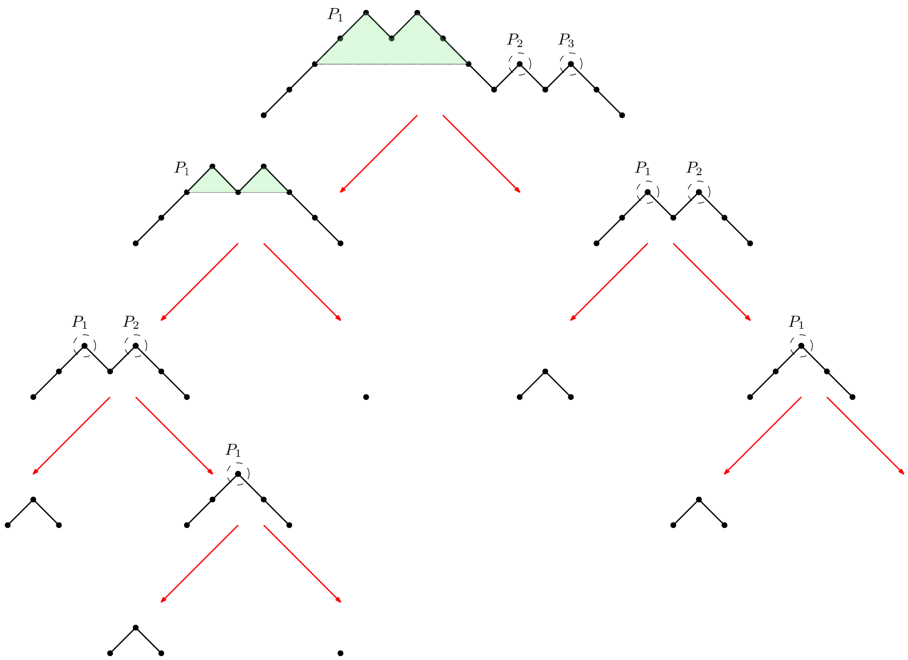



FIGURE 10. Example of the bijection $\nu$.

Next, by combining the last equations of (1) and (2), we eliminate the variable $D_{\{1\}}(x,y)$. This yields the following polynomial equation:

$$\begin{aligned}-1+D_{\{2,3\}}(x,y)-x+C(x)-C(x)D_{\{2,3\}}(x,y)x-x^2+C(x)x^2+2xy\\ -2D_{\{2,3\}}(x,y)xy+2x^2y-C(x)x^2y\\ -D_{\{2,3\}}(x,y)x^2y+C(x)D_{\{2,3\}}(x,y)x^2y-x^2y^2+D_{\{2,3\}}(x,y)x^2y^2=0.\end{aligned}$$

Solving this equation yields the desired generating function $D_{\{2,3\}}(x,y)$. For any given finite set $S\subset\mathbb{Z}^+$, our implementation computes a polynomial

$$P(x,y,C(x),D_S(x,y)).$$

This procedure is implemented in *Mathematica* via the function `PolDS`$[S]$. For example, for the set $S=\{2,3\}$, we obtain:

In[1]:= `PolDS[{2, 3}]`

Out[1]= $-1 - x + C(x)\ x - x^2 + C(x)\ x^2 + 2\ x\ y + 2\ x^2\ y - C(x)\ x^2\ y - x^2\ y^2 + D_S - C(x)\ x\ D_S - 2\ x\ y\ D_S - x^2\ y\ D_S + C(x)\ x^2\ y\ D_S + x^2\ y^2\ D_S$

To express the solution in terms of the generating function of the Catalan numbers, we use the command `PolDSCat`$[S]$. For our example, this gives:

In[2]:= `PolDSCat[{2, 3}]`

Out[2]= $\dfrac{1 + x - C(x)\ x + x^2 - C(x)\ x^2 - 2\ x\ y - 2\ x^2\ y + C(x)\ x^2\ y + x^2\ y^2}{1 - C(x)\ x - 2\ x\ y - x^2\ y + C(x)\ x^2\ y + x^2\ y^2}$

For the explicit generating function, we use the command `GeneratingFunctionDS`$[S, x, y]$. For instance:

```
In[3]:= GeneratingFunctionDS[{2, 3}, x, y]//Factor
```

$$\text{Out[3]=}\ \frac{-1 - \sqrt{1-4x} - x - \sqrt{1-4x}\ x - 2x^2 + 3xy + \sqrt{1-4x}\ xy + 4x^2y - 2x^2y^2}{-1 - \sqrt{1-4x} + 3xy + \sqrt{1-4x}\ xy + 2\ x^2y - 2x^2y^2}$$

Finally, to compute the first $n$ rows and columns of the array $\mathcal{N}_S^{(1)}$, we use the command `NarayanaMatrixT1[S, n]`. For our example, this gives:

```
In[4]:= NarayanaMatrixT1[{2, 3}, 8]
```

Out[4]=

$$\begin{pmatrix} 1 & 0 & 0 & 0 & 0 & 0 & 0 \\ 1 & 0 & 0 & 0 & 0 & 0 & 0 \\ 1 & 1 & 0 & 0 & 0 & 0 & 0 \\ 1 & 3 & 1 & 0 & 0 & 0 & 0 \\ 2 & 5 & 6 & 1 & 0 & 0 & 0 \\ 5 & 11 & 15 & 10 & 1 & 0 & 0 \\ 14 & 29 & 38 & 35 & 15 & 1 & 0 \\ 42 & 84 & 107 & 104 & 70 & 21 & 1 \end{pmatrix}$$

## 8. Acknowledgments

The first author was partially supported by ANR PiCs (ANR-22-CE48-0002). The second author was partially supported by the Citadel Foundation, Charleston, SC. The last author was partially supported by Universidad Nacional de Colombia, Project No. 64041.

## References

[1] E. Barcucci, A. Bernini, and R. Pinzani, Exhaustive generation of some lattice paths and their prefixes, *Theoret. Comput. Sci.* **478** (2021), 47–52.

[2] J.-L. Baril and S. Kirgizov, Pure descent statistic on permutations, *Discrete Math.* **340** (2017), 2550–2558.

[3] J.-L. Baril and A. Petrossian, Equivalence classes of Dyck paths modulo some statistics, *Discrete Math.* **338** (2015), 655–660.

[4] J. J. Bravo, J. L. Herrera, and J. L. Ramírez, Combinatorial Interpretation of Generalized Pell Numbers, *J. Integer Seq.* **23** (2020), Article 20.2.1

[5] M. Dairyko, L. Pudwell, S. Tyner, and C. Wynn, Non-contiguous pattern avoidance in binary trees, *Electron. J. Combin.* **19** (2012), # P22.

[6] F. R. Bernhart, Catalan, Motzkin, and Riordan numbers, *Discrete Math.* **204** (1999), 73–112.

[7] AJ Bu, Automated counting of restricted Motzkin paths, *Enumer. Comb. Appl.* **1** (2021), Article #S2R12.

[8] AJ Bu and D. Zeilberger, Using symbolic computation to explore generalized Dyck paths and their areas, *Palest. J. Math.* **13** (2024), 1–13.

[9] E. Deutsch, Dyck path enumeration, *Discrete Math.* **204** (1999), 167–202.

[10] S. B. Ekhad and D. Zeilberger, Automatic counting of restricted Dyck paths via (numeric and symbolic) dynamic programming. Preprint, arXiv:2006.01961, (2020).

[11] S.-P. Eu, S.-C. Liu, and Y.-N. Yeh, Dyck paths with peaks avoiding or restricted to a given set, *Stud. Appl. Math.* **111** (2003), 453–465.

[12] P. Flajolet and R. Sedgewick, *Analytic Combinatorics*, Cambridge University Press, Cambridge, 2009.

[13] R. Flórez and J. L. Ramírez, Enumerating symmetric and asymmetric peaks in Dyck paths, *Discrete Math.* **343** (2020), 112118.
[14] T. Mansour, Statistics on Dyck paths, *J. Integer Sequences.* **9** (2006), Article 06.1.5.
[15] D. Merlini, R. Sprugnoli, and M.C. Verri, Some statistics on Dyck paths, *J. Statis. Plann. Inference* **101** (2002), 211–227.
[16] T. K. Petersen, *Eulerian Numbers*, Birkhäuser Advanced Texts Basler Lehrbücher, 2015.
[17] A. Sapounakis, I. Tasoulas, and P. Tsikouras, Counting strings in Dyck paths, *Discrete Math.* **307** (2007), 2909–2924.
[18] N. J. A. Sloane, The On-Line Encyclopedia of Integer Sequences. http://oeis.org/.
[19] R. Stanley, *Catalan Numbers*, Cambridge University Press, 2015.

LIB, Université de Bourgogne, B.P. 47 870, 21078, Dijon Cedex, France
*Email address*: barjl@u-bourgogne.fr

Department of Mathematical Sciences, The Citadel, Charleston, SC, U.S.A.
*Email address*: rflorez1@citadel.edu

Departamento de Matemáticas, Universidad Nacional de Colombia, Bogotá, Colombia
*Email address*: jlramirezr@unal.edu.co